REMARKS

Reconsideration of the present application is respectfully requested. Claims 5, 15 and 21 have been canceled. Claims 1, 4, 6, 7, 11, 16, 17, 22, 23 and 26 have been amended. No new matter has been added.

In the Office Action, claims 1-28 were rejected under 35 U.S.C. § 102(e) based on U.S. Patent no. 6,539,352 of Sharma et al. ("Sharma"). Applicants respectfully traverse the rejections. The amendments to the claims are made only to more clearly recite what Applicants regard as the invention, not in response to the rejections or to comply with any statutory requirement of patentability,.

The present invention relates to an integrated speaker and speech recognition system that provides improved speaker-specific response in a noisy environment. In certain embodiments, the invention combines the results of automatic speaker verification and automatic speech recognition to select one of multiple automatic speech recognition hypotheses. For example, claim 23 recites:

> 23. (Currently amended) A method comprising:
> receiving an utterance from an intended talker at a speech recognition
>   system;
> computing a speaker verification score based on a voice characteristic model
>   associated and with the utterance;
> computing a speech recognition score associated with the utterance; and
> **selecting a best hypothesis from a plurality of hypotheses representing
>   automatic speech recognition results of the utterance, based on both
>   the speaker verification score and the speech recognition score.**
>   (Emphasis added.)

10

Before discussing this further, one must recognize the difference between recognizing speech ("speech recognition") and recognizing (verifying) a speaker ("speaker verification"). Speech recognition involves recognizing what a speaker is saying. Speaker verification, on the other hand, involves determining whether the speaker is who he claims to be. This distinction is well-understood by those skilled in the relevant art. With this in mind, consider Sharma.

Sharma does not disclose or even suggest a method such as recited in claim 23. In particular, Sharma does not disclose computing a speaker verification score and a speech recognition score based on an utterance. Further, Sharma does not disclose that a speaker verification score and a speech recognition score can be used to select one of a plurality of hypotheses representing automatic speech recognition results of the utterance as a best hypothesis.

Regarding the first point (the use of speaker verification and speech recognition scores), the Examiner cites Sharma at col. 4, lines 33-34; col. 3, lines 41-43; col. 4 line 28; col. 4 lines 27-31; and col. 5 lines 10-11 (see Office Action p. 3). However, the cited disclosure in Sharma does not disclose or suggest the above-mentioned claim features, nor are such features found anywhere else in Sharma. The entire system of Sharma is a speaker verification system. The ultimate output of the Sharma system, therefore, is essentially a yes/no decision on whether the speaker is who he claims to be. Sharma discloses that during the verification process:

> The multiple classifiers of the enrollment component are used to 'score' the subword data, and the scores are fused, or combined. The result of the fusion is a "final score". The final score is compared to the stored threshold. If the final score exceeds the threshold, the test sample

is verified as the user's. If the final score is less than the threshold, the test sample is declared not to be the user's. Sharma, col. 5, lines10-17.

Sharma further states:

In the preferred embodiment, a classifier fusion module 130 using the linear opinion pool method combines the NTN score and the GMM score. Col. 11, lines 43-45.

The threshold value output 140 is compared to a 'final score' in the testing component to determine whether a test user's voice has so closely matched the model that it can be said that the two voices are from the same person. Col. 12, lines 8-12.

Thus, the "scores" mentioned in Sharma merely indicate how closely segments of speech match a model of previously stored speech. If the Examiner considers these scores to be "speaker verification scores" as recited Applicant's claims, then they cannot also be considered to be speech recognition scores per Applicant's claims. Likewise, if the Examiner considers these scores to be "speech recognition scores" as recited Applicant's claims, then they cannot also be considered to be speaker verification scores per Applicant's claims. Sharma clearly dos not disclose or suggest computing a speaker verification score and a speech recognition score based on a given utterance.

Second, even assuming *arguendo* Sharma does disclose computing a speaker verification score and a speech recognition score based on an utterance, Sharma still does not disclose that both types of scores can be used to select a hypothesis representing an automatic speech recognition result of the utterance as a best hypothesis. The Examiner contends that the selection of a best hypothesis is

"inherently" disclosed at col. 5, lines 14-16 of Sharma. However, that is not true. Sharma only discloses comparing the (fused) scores to a threshold to make a yes/no decision regarding whether the test sample (voice) is the user's. The scores are not used to select any hypothesis, much less to select a speech recognition hypothesis, and certainly not to select one of a plurality of speech recognition hypotheses as a best hypothesis.

Furthermore, the Examiner's reliance upon inherency is improper. "Inherency . . . may not be established by mere probabilities or possibilities. The mere fact that a certain thing may result from a given set of circumstances is not sufficient." Continental Can Co. v. Monsanto Co., 948 F.2d 1264, 1269, 20 U.S.P.Q.2d 1746, 1749 (Fed. Cir. 1991) (quoting In re Oelrich, 666 F.2d 578, 581, 212 U.S.P.Q. 323, 326 (C.C.P.A. 1981). If the Examiner intends to maintain the rejection using this rationale, the Examiner must provide evidence that this claim feature is in fact inherent in the system of Sharma, in the same manner recited in Applicant's claims.

Hence, Sharma does not disclose or suggest computing a speaker verification score and a speech recognition score based on an utterance, and using both scores to select one of a plurality of hypotheses representing automatic speech recognition results of the utterance as a best hypothesis. For at least these reasons, therefore, claim 23 and all claims which depend on it patentable over the cited art.

Each of the remaining independent claims also includes limitations similar to those discussed above regarding claim 23. Therefore, all of Applicant's claims are patentable over the cited art, for at least these reasons.

13

Dependent Claims

In view of the above remarks, a specific discussion of the dependent claims is considered to be unnecessary. Therefore, Applicants' silence regarding any dependent claim is not to be interpreted as agreement with, or acquiescence to, the rejection of such claim or as waiving any argument regarding that claim.
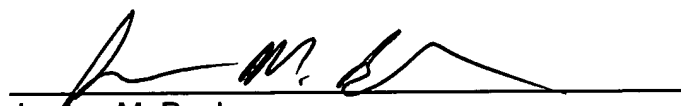
Conclusion

For the foregoing reasons, the present application is believed to be in condition for allowance, and such action is earnestly requested.

If any additional fee is required, please charge Deposit Account No. 02-2666.

Respectfully submitted,
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP

Date: 8/25/05

Jordan M. Becker
Reg. No. 39,602

Customer No. 08791
12400 Wilshire Boulevard
Seventh Floor
Los Angeles, CA 90025-1030
(408) 720-8300

14